

# SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Rubin, Gerald M.

Pan, Duoia  
Rooke, Jenny  
Yavari, Reza  
Xu, Tian

(ii) TITLE OF INVENTION: KUZ: A Novel Family of Metalloproteases

(iii) NUMBER OF SEQUENCES: 8

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
(B) STREET: 268 BUSH STREET, SUITE 3200  
(C) CITY: SAN FRANCISCO  
(D) STATE: CALIFORNIA  
(E) COUNTRY: USA  
(F) ZIP: 94104

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: OSMAN, RICHARD A  
(B) REGISTRATION NUMBER: 36,627  
(C) REFERENCE/DOCKET NUMBER: B97-081

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (415) 343-4341  
(B) TELEFAX: (415) 343-4342

## (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5630 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTTTAAAAAA	AACCACCAAG	CGAGTTGGAC	GCGTAACTCT	TTGTAACGGA	TCTCGGAACG	60
CCGTGGGAGT	CGGAAAATCG	CTGGACGCGT	GTTTCGTGCGT	TTGCATGTGT	GCGTGCGTTC	120
GTGTGTGTGT	GTGTGCTAAT	GTGCGAGCGG	GTGAGCGAAT	AAAAATAAAT	ATATATCGTC	180
AAGTCAGGCT	TAAGAAATGT	GCGCTAATCA	AAGAAAATGC	CCCCAATTCT	GGCCAATTGA	240
GAATTGTGGC	TAAACAAAAA	ATTCGACCGG	AGTTCAAAAA	TAAACAATCC	AGTGAATAAA	300
CACACAAAAAT	CAATCAAAAA	AGAAGATTTT	TCTTTTTTAT	TTTCGCTTTT	AATTTATTAA	360
CGAGAATAAT	AAATAAATAA	ATAAATAAAT	ATAAACAAAA	ATAAAAATAT	AAGAAAAGTG	420
TACGTGACAA	GAGCTCGAAA	AGAAGTTGCA	ACAAATAGCA	AAAATAATTC	GTGCGTGCGA	480
AAAAGTGCTG	CGAAGTTTTA	TGGCCCATGC	AAAAAGTGCT	AAATTTGTAA	ATGGCATGGA	540
AAGTGCAAAG	CTCTGATTAA	AAAACCCGCG	AAGATTGGAG	TGCGAGGTGC	CGCCCAATAA	600
CGCAACCAAC	TACTGCCACA	AGGAAATTAT	TAAGACCAAT	CAACGACCAA	AAAAATAAAA	660
AATAAAACAA	AAGCAAGCAG	AAATTTGGTG	CTAGTTCTGT	TTAGTCGACA	GCCATCCACG	720
TTGGATCCCC	ATCGCAAATA	ATGTCATCAA	AATGTGCTTT	CAACATTGTA	TTCGTATCGA	780

	TCATTTTTCAT	CATCATCGTA	AATGGTTACG	CAAAAGATAT	TTCTGGAGTT	AAAAGAGGTC	840
	ATGAACGACT	TAACGAATAC	ATATCCCAC	ATGAAACACT	CAACTATGAT	CACGAGCACA	900
	TCCGAGCTAG	TCACAATAGA	GCGCGACGAT	CAGTGACCAA	AGATCAATAT	GTACATTTAA	960
5	AGTTTGCATC	ACATGGAAGA	GACTTCCATC	TTAGATTAAA	ACGTGATTTA	AATACATTTA	1020
	GCAATAAGTT	AGACTTTTAT	GATAGCAAAG	GTCCCATTGA	TGTCTCCACG	GATCATATCT	1080
	ATGAGGGCGA	AGTGATAGGG	GATCGTAATA	GTTATGTATT	TGGTTCCATA	CACAATGGGG	1140
	TATTCGAGGG	TAAAATTATA	ACGGAACGTG	ATGCCTATTA	TGTTGAACAT	GCCAAACATT	1200
	ATTTTCCCAC	AAATCGCACG	GCGACAACAA	CACCACCATC	GACTTCGACG	ACATCCTCAG	1260
10	CAACAACAGT	CACAAAAAGC	ACACAACCAA	CACGGCCTTT	GGCCAAAAGC	AACACCAGTA	1320
	CTACTGCCGT	TAATAGTAAG	ACAGAAAAC	TTATAAAGAA	AATTGCTGAA	TCCACAACGA	1380
	CGAGCCAGCA	GCTTCCAGAA	TATACCGAAT	CGTCGTCGTC	GTCGTCGACA	ACAACATTCC	1440
	CACCCACAAC	AGAGTATTTT	GAGGACGAAA	AGGAGCGTAA	TGCCGAGGAC	GAACCTTGATT	1500
	TTCACCTCCAT	TATCTACAAG	GAGTCACATG	TCGAGGACGC	CTACGAAAAT	GTGCGCGAAG	1560
	ATTCACGTGGC	CGGCTGTGGC	ATCACGGATG	AGGTCTCTCA	GTGGATGGAG	AACATACAAA	1620
15	AGTCAGCCGT	CGAAGAGTTG	CCGGAGCCCA	TGTCAAAGGA	CTATCAAAG	CTCCACCGGA	1680
	AGCAGCTGCA	CAAAAAGTCC	GCCCCACAGC	AACAACAGCA	GCCCCATCCG	CCGAAGAAGT	1740
	ACATCAGCGG	GGATGAGGAC	TTCAAGTATC	CCCACCAGAA	GTACACGAAG	GAAGCTAACT	1800
	TCGCCGAGGG	TGCATTCTAC	GATCCATCGA	CCGGACGTCG	CCTGGGCTCA	TCCGCCAACG	1860
	TGGCCGACTG	GCATCAGCTC	GTCCACGAGC	GCGTCCGCCG	CGCCACCGAC	AATGGTGCTG	1920
20	GGGATAGGGG	CTCATCCGGT	GGATCTGGAC	GCGGTCGCGA	GGACAACAAG	AATACCTGCT	1980
	CGCTCTACAT	TCAAACGGAT	CCATTGATAT	GGCGCCACAT	ACGCGAAGGC	ATTGCTGACC	2040
	ACGATCGTGG	ACGCAAGTAC	GAGGTGGATG	AGAAAACGCG	CGAGGAAATC	ACATCGTTGA	2100
	TTGCACATCA	CGTGACGGCC	GTTAATTACA	TTTACCGCAA	CACAAAGTTC	GACGGACGCA	2160
	CCGAGCATCG	CAACATACGC	TTTGAGGTGC	AACGCATTAA	GATCGATGAC	GATTCCGGCT	2220
25	GTGCAATTC	CTACAATGGT	CCACACAATG	CCTTTTGTCAA	TGAACACATG	GATGTCTCGA	2280
	ACTTTTTTGAA	TCTGCATTCC	CTAGAAGATC	ACTCGGACTT	TTGTTTGGCT	TACGTGTTCA	2340
	CCTACAGAGA	TTTCACTGGC	GGCACTTTGG	GTCTGGCCTG	GGTGGCCAGT	GCGTCGGGAG	2400
	CCTCTGGTGG	AATTTGCGAG	AAGTACAAGA	CGTACACGGA	AACGGTGGGT	GGACAGTACC	2460
	AGAGCACCAA	GCGATCACTC	AACACGGGCA	TCATCACCTT	TGTCAACTAC	AACAGTCGGG	2520
30	TGCCGCCGAA	AGTGTGCGAG	CTTACGTTGG	CACACGAGAT	TGGCCACAAC	TTTGGATCAC	2580
	CTCACGATTA	CCCTCAGGAA	TGTCGTCTCTG	GTGGCCATAA	TGGCAATTAC	ATTATGTTCC	2640
	CCAGTGCCAC	CTCCGGTGAT	AGGCCAAATA	ACTCCAAGTT	CTCGCCCTGC	TCCATTCCGA	2700
	ACATCTCCAA	TGTCTTGAG	GTGCTGGTGG	GCAACACGAA	GCGCGACTGC	TTCAAGGCCT	2760
	CGGAAGGTGC	CTTCTGCGGC	AACAAGATCG	TGGAGTCTGG	CGAGGAATGC	GAAGTGGCT	2820
35	TCAACGAGGA	GGAGTGCAAG	GACAAGTGCT	GCTACCCGCG	TCTGATCAGC	GAGTACGACC	2880
	AGTCGCTGAA	CTCCAGTGCC	AAGGGATGCA	CGCGCCGCGC	CAAGACCCAG	TGCTCACCAT	2940
	CGCAGGGTCC	GTGCTGTCTG	TCCAACCTCT	GCACCTTTGT	GCCGACGAGC	TACCACCAGA	3000
	AGTGCAAGGA	GGAGACGGAG	TGCAGCTGGT	CGAGCACATG	CAACGGAACC	ACGGCCGAGT	3060
	GTCCGGAGCC	ACGTATCATG	GATGACAAGA	CCATGTGCAA	CAATGGAACA	GCGCTATGCA	3120
40	TCCGCGGTGA	ATGTAGTGGA	TCGCCATGTT	TGCTCTGGAA	TATGACAAAG	TGCTTCCTTA	3180
	CCTCGACCAC	ACTGCCGCAC	GTGAGCAAGC	GCAAGTTGTG	CGACTTGGCC	TGCCAGGATG	3240
	GCAATGACAC	CTCCACCTGC	CGCAGCACCA	GCGAGTTTGC	CGATAAATAT	AATATTCAAA	3300
	AGGGTGGTAT	TAGTCTGCAG	CCCGGTTTCG	CATGCGATAA	TTTCCAGGGC	TACTGCGATG	3360
	TGTTCCCTTAA	GTGTCGAGCC	GTGGATGCCG	ATGGTCCGCT	TCTTCGGCTG	AAGAATTTGT	3420
45	TGCTCAACCG	GAAGACCCTG	CAAACGGTGG	CCGAGTGGAT	CGTCGACAAT	TGGTACCTAG	3480
	TGGTTCTGAT	GGGAGTGGCC	TTTATTGTGG	TCATGGGTTT	GTTTCATCAA	TGTTGTGCCG	3540
	TGCACACGCC	CAGTTCCAAT	CCGAAGAAGC	GACGAGCTCG	TCGAATCAGC	GAAACTCTAA	3600
	GAGCACCCAT	GAACACGTTG	CGTAGAATGC	AACGTCATCC	CAATCAGCGA	GGAGCAGGTC	3660
	CTCGAAGCAT	CCCACCGCCG	GCACATGAGG	CGCAGCATT	TTCACGCGGC	GGAGATGGTC	3720
50	GCGGCGGCGG	CGGTGGAGGC	GGAGGTCGCC	ACGGTGGCTC	TAGGTCACAC	CATCAACAGC	3780
	ATCCGCACGA	TTGGGATCGT	CATCAGGGTG	GCCACTCAAT	CGTCCCATTG	CCCACCGGCG	3840
	GCAGCCATTC	AAGTCGCAAC	TCGGCGGCGA	ATCAAGCGAG	AAGAAGCGAT	GGACGAGGTC	3900
	CACGATCCAC	CAGCAGTGGG	CGGCCGCGAG	CTATAGCCAG	CGGAAGCGGT	GCCGCGAGCG	3960
	GAGCAGCGCG	ATCTCATGGC	GGGTACGGAG	CCGAACAGGC	GATACCGGGT	TCCATTGGTG	4020
55	GTGGTGTCCA	GGCGGCCATT	AGCAGCGGCG	GTGTGGTGGC	TCGGGCCCG	CTGCCGCTGC	4080

5 CATTGCCGCC GCCAAATGGA CAGCAGCAAA TGCAACAGCA ACAACAACCTG CAACTACAGC 4140  
 AACCGGCAAT TTCGCCGCAG CAGCAGCCGC AGCAAGCGTT CTACACGCCG AAAGAAGTAC 4200  
 CACCACGCAA TAAGTCCCGA TCATCACGTA CCAACAACAC CTCCAACACC ACAACCACCA 4260  
 CCAACTCATC CACAGCGGCA GCCGGCAGTG GGTCGGTCTC GGGACCGGGC TCGGGGGCGG 4320  
 GCAGTAGTAG TAAGAGCAAG AGCGGTAAAA GTGCCAAAGC CAAAGACTCA AAGTCGCAAA 4380  
 AATCGCAGCA GGCCAACAAC AGTCGCAGCA GCAGCAAGGA GAAGGGCGTC AAGCCAGTGC 4440  
 GCCGAAATAT CGTTTATTAG GAGCGGAACC ATCACATTGC CATAACAAC ACTGAACGAA 4500  
 ATATAGCCCC GAACCCAAAA TATCAAATGC AACCACATAT AGAATCGCCC GCTGCTAGTC 4560  
 ATCGAACTAC ATGTATGAGT TGTGTCTTCC CATCCACCGA CAAACACAAA CAGAAAAGAA 4620  
 10 ATTATAATGA TATTTCATTT AATCGATGCA ATTGGCGTCG CGCCGCCTCC GCTACAAGTA 4680  
 AGCTTTAGTC GGCCGACATC GTTGACGAG CAACAGCAGC AGCAACATCA TCTGCAGCAG 4740  
 CAGCAGCAGC ATCAGCAGCA ACTGGAGCCG CAGCAGCAAC ACGCCTATGC CGATGCTTAT 4800  
 GCGGCCTTGG GGCGGGGCCA GTATGAGTCC ACCACGCGGG CGCCCAACAA CAGCAAGGTT 4860  
 TGACAGCCAA AAGTAGCAAT GGAGCGCCAC AAAAGGCCAA AGGCTAAGCG ACTCAAGCAG 4920  
 15 CAGAAGGAGC CGCATACACA GCAAACAACA ACACAGCAAC AAAAGCAAAA ACAACATAAA 4980  
 TCAAATGAAC TCAAATTAATA TGTAATGTAT ATTTTATATG TAATTATTTT TATTTAAACA 5040  
 GTGTTTGTAT GCCACAAGGG AAAACAGCCA GCAACAAAAA GAAAAATACA AAAATAACAC 5100  
 AAAAAAGGAG ACAAATTTTCG TAATACAGAA AAAGCTGAAA GTGAATGATA TTTTGTATTA 5160  
 ACTAAATTA AATGAAAATA CGAATGCAAA TTATGAATAA TAAAAGTAAT TAAAAACGAC 5220  
 20 AACATGCATA ATACATATAA AGTTGCAAGT TGCATATATA TACATTTGTA TGTATATATT 5280  
 TATTATGGAT ACACAATTAT TAAATAGCAG CAGCCACAAC AAACAAGTAA TATACATGAA 5340  
 GAAAAACTAA GGTTTAATTG TATGAGAAAG CATTCTATAT GTCGGTGAGA TTTCTAAGCG 5400  
 CTAGGCCGAA ATACAAAATT AATTACACAC TTGAATAACA AAATGTGTTT TGTACAAAAA 5460  
 AAAAAAATG AAATAAACAA AAACAGTGCG AATTAATTAA GCGTCATTAT AAAAAAAGA 5520  
 25 ACGGAAACAA CAAAGCATTT AAATTGTATT TATCTGTACC GAAGCTAAAC GTTTATTTAA 5580  
 AGCCGTCAAA ATTGCAATTTG TAACTAGCA AAACAAAAAA AAAAAAAAC 5630

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1239 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ser Ser Lys Cys Ala Phe Asn Ile Val Phe Val Ser Ile Ile Phe  
 1 5 10 15  
 Ile Ile Ile Val Asn Gly Tyr Ala Lys Asp Ile Ser Gly Val Lys Arg  
 20 25 30  
 40 Gly His Glu Arg Leu Asn Glu Tyr Ile Ser His Tyr Glu Thr Leu Asn  
 35 40 45  
 Tyr Asp His Glu His Ile Arg Ala Ser His Asn Arg Ala Arg Arg Ser  
 50 55 60  
 45 Val Thr Lys Asp Gln Tyr Val His Leu Lys Phe Ala Ser His Gly Arg  
 65 70 75 80  
 Asp Phe His Leu Arg Leu Lys Arg Asp Leu Asn Thr Phe Ser Asn Lys  
 85 90 95  
 Leu Asp Phe Tyr Asp Ser Lys Gly Pro Ile Asp Val Ser Thr Asp His  
 100 105 110  
 50 Ile Tyr Glu Gly Glu Val Ile Gly Asp Arg Asn Ser Tyr Val Phe Gly  
 115 120 125  
 Ser Ile His Asn Gly Val Phe Glu Gly Lys Ile Ile Thr Glu Arg Asp  
 130 135 140  
 55 Ala Tyr Tyr Val Glu His Ala Lys His Tyr Phe Pro Thr Asn Arg Thr  
 145 150 155 160

Ala Thr Thr Thr Pro Pro Ser Thr Ser Thr Thr Ser Ser Ala Thr Thr  
165 170 175  
Val Thr Lys Ser Thr Gln Pro Thr Arg Pro Leu Ala Lys Ser Asn Thr  
180 185 190  
5 Ser Thr Thr Ala Val Asn Ser Lys Thr Glu Asn Phe Ile Lys Lys Ile  
195 200 205  
Ala Glu Ser Thr Thr Thr Ser Gln Gln Leu Pro Glu Tyr Thr Glu Ser  
210 215 220  
Ser Ser Ser Ser Ser Thr Thr Thr Phe Pro Pro Thr Thr Glu Tyr Phe  
225 230 235 240  
10 Glu Asp Glu Lys Glu Arg Asn Ala Glu Asp Glu Leu Asp Phe His Ser  
245 250 255  
Ile Ile Tyr Lys Glu Ser His Val Glu Asp Ala Tyr Glu Asn Val Arg  
260 265 270  
15 Glu Gly His Val Ala Gly Cys Gly Ile Thr Asp Glu Val Ser Gln Trp  
275 280 285  
Met Glu Asn Ile Gln Asn Ser Ala Val Glu Glu Leu Pro Glu Pro Met  
290 295 300  
20 Ser Lys Asp Tyr Gln Lys Leu His Arg Lys Gln Leu His Lys Lys Ser  
305 310 315 320  
Ala Pro Gln Gln Gln Gln Gln Pro His Pro Pro Lys Lys Tyr Ile Ser  
325 330 335  
Gly Asp Glu Asp Phe Lys Tyr Pro His Gln Lys Tyr Thr Lys Glu Ala  
340 345 350  
25 Asn Phe Ala Glu Gly Ala Phe Tyr Asp Pro Ser Thr Gly Arg Arg Leu  
355 360 365  
Gly Ser Ser Ala Asn Val Ala Asp Trp His Gln Leu Val His Glu Arg  
370 375 380  
30 Val Arg Arg Ala Thr Asp Asn Gly Ala Gly Asp Arg Gly Ser Ser Gly  
385 390 395 400  
Gly Ser Gly Arg Gly Arg Glu Asp Asn Lys Asn Thr Cys Ser Leu Tyr  
405 410 415  
Ile Gln Thr Asp Pro Leu Ile Trp Arg His Ile Arg Glu Gly Ile Ala  
420 425 430  
35 Asp His Asp Arg Gly Arg Lys Tyr Glu Val Asp Glu Lys Thr Arg Glu  
435 440 445  
Glu Ile Thr Ser Leu Ile Ala His His Val Thr Ala Val Asn Tyr Ile  
450 455 460  
40 Tyr Arg Asn Thr Lys Phe Asp Gly Arg Thr Glu His Arg Asn Ile Arg  
465 470 475 480  
Phe Glu Val Gln Arg Ile Lys Ile Asp Asp Asp Ser Ala Cys Arg Asn  
485 490 495  
Ser Tyr Asn Gly Pro His Asn Ala Phe Cys Asn Glu His Met Asp Val  
500 505 510  
45 Ser Asn Phe Leu Asn Leu His Ser Leu Glu Asp His Ser Asp Phe Cys  
515 520 525  
Leu Ala Tyr Val Phe Thr Tyr Arg Asp Phe Thr Gly Gly Thr Leu Gly  
530 535 540  
50 Leu Ala Trp Val Ala Ser Ala Ser Gly Ala Ser Gly Gly Ile Cys Glu  
545 550 555 560  
Lys Tyr Lys Thr Tyr Thr Glu Thr Val Gly Gly Gln Tyr Gln Ser Thr  
565 570 575  
Lys Arg Ser Leu Asn Thr Gly Ile Ile Thr Phe Val Asn Tyr Asn Ser  
580 585 590  
55 Arg Val Pro Pro Lys Val Ser Gln Leu Thr Leu Ala His Glu Ile Gly

595 600 605

His Asn Phe Gly Ser Pro His Asp Tyr Pro Gln Glu Cys Arg Pro Gly

610 615 620

Gly Leu Asn Gly Asn Tyr Ile Met Phe Ala Ser Ala Thr Ser Gly Asp

625 630 635 640

Arg Pro Asn Asn Ser Lys Phe Ser Pro Cys Ser Ile Arg Asn Ile Ser

645 650 655

Asn Val Leu Asp Val Leu Val Gly Asn Thr Lys Arg Asp Cys Phe Lys

660 665 670

Ala Ser Glu Gly Ala Phe Cys Gly Asn Lys Ile Val Glu Ser Gly Glu

675 680 685

Glu Cys Asp Cys Gly Phe Asn Glu Glu Glu Cys Lys Asp Lys Cys Cys

690 695 700

Tyr Pro Arg Leu Ile Ser Glu Tyr Asp Gln Ser Leu Asn Ser Ser Ala

705 710 715 720

Lys Gly Cys Thr Arg Arg Ala Lys Thr Gln Cys Ser Pro Ser Gln Gly

725 730 735

Pro Cys Cys Leu Ser Asn Ser Cys Thr Phe Val Pro Thr Ser Tyr His

740 745 750

Gln Lys Cys Lys Glu Glu Thr Glu Cys Ser Trp Ser Ser Thr Cys Asn

755 760 765

Gly Thr Thr Ala Glu Cys Pro Glu Pro Arg His Arg Asp Asp Lys Thr

770 775 780

Met Cys Asn Asn Gly Thr Ala Leu Cys Ile Arg Gly Glu Cys Ser Gly

785 790 795 800

Ser Pro Cys Leu Leu Trp Asn Met Thr Lys Cys Phe Leu Thr Ser Thr

805 810 815

Thr Leu Pro His Val Ser Lys Arg Lys Leu Cys Asp Leu Ala Cys Gln

820 825 830

Asp Gly Asn Asp Thr Ser Thr Cys Arg Ser Thr Ser Glu Phe Ala Asp

835 840 845

Lys Tyr Asn Ile Gln Lys Gly Gly Ile Ser Leu Gln Pro Gly Ser Pro

850 855 860

Cys Asp Asn Phe Gln Gly Tyr Cys Asp Val Phe Leu Lys Cys Arg Ala

865 870 875 880

Val Asp Ala Asp Gly Pro Leu Leu Arg Leu Lys Asn Leu Leu Leu Asn

885 890 895

Arg Lys Thr Leu Gln Thr Val Ala Glu Trp Ile Val Asp Asn Trp Tyr

900 905 910

Leu Val Val Leu Met Gly Val Ala Phe Ile Val Val Met Gly Ser Phe

915 920 925

Ile Lys Cys Cys Ala Val His Thr Pro Ser Ser Asn Pro Lys Lys Arg

930 935 940

Arg Ala Arg Arg Ile Ser Glu Thr Leu Arg Ala Pro Met Asn Thr Leu

945 950 955 960

Arg Arg Met Gln Arg His Pro Asn Gln Arg Gly Ala Gly Pro Arg Ser

965 970 975

Ile Pro Pro Pro Ala His Glu Ala Gln His Tyr Ser Arg Gly Gly Asp

980 985 990

Gly Arg Gly Gly Gly Gly Gly Gly Gly Arg His Gly Gly Ser Arg

995 1000 1005

Ser His His Gln Gln His Pro His Asp Trp Asp Arg His Gln Gly Gly

1010 1015 1020

His Ser Ile Val Pro Leu Pro Thr Gly Gly Ser His Ser Ser Arg Asn

1025 1030 1035 1040

Ser Ala Ala Asn Gln Ala Arg Arg Ser Asp Gly Arg Gly Pro Arg Ser  
 1045 1050 1055  
 Thr Ser Ser Gly Arg Pro Gln Ala Ile Ala Ser Gly Ser Gly Ala Ala  
 1060 1065 1070  
 5 Ser Gly Ala Ala Arg Ser His Gly Gly Tyr Gly Ala Glu Gln Ala Ile  
 1075 1080 1085  
 Pro Gly Ser Ile Gly Gly Gly Val Gln Ala Ala Ile Ser Ser Gly Gly  
 1090 1095 1100  
 10 Val Val Ala Arg Ala Gln Leu Pro Leu Pro Leu Pro Pro Pro Asn Gly  
 1105 1110 1115 1120  
 Gln Gln Gln Met Gln Gln Gln Gln Gln Leu Gln Leu Gln Gln Pro Ala  
 1125 1130 1135  
 Ile Ser Pro Gln Gln Gln Pro Gln Gln Ala Phe Tyr Thr Pro Lys Glu  
 1140 1145 1150  
 15 Leu Pro Pro Arg Asn Lys Ser Arg Ser Ser Arg Thr Asn Asn Thr Ser  
 1155 1160 1165  
 Asn Thr Thr Thr Thr Thr Asn Ser Ser Thr Ala Ala Ala Gly Ser Gly  
 1170 1175 1180  
 20 Ser Val Ser Gly Pro Gly Ser Gly Ala Gly Ser Ser Ser Lys Ser Lys  
 1185 1190 1195 1200  
 Ser Gly Lys Ser Ala Lys Ala Lys Asp Ser Lys Ser Gln Lys Ser Gln  
 1205 1210 1215  
 Gln Ala Asn Asn Ser Arg Ser Ser Ser Lys Glu Lys Gly Val Lys Pro  
 1220 1225 1230  
 25 Val Arg Arg Asn Ile Val Tyr  
 1235

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2796 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAATTCGCGG TTTTGGAGGA GCTAGGAGCG TTGCCGCGCC CTGAAGTGGA GCGAGAGGGA 60  
 GGTGCTTTTC CCGTTCTCCT GCCAGGGGAG GTCCCGGCTT CCCGTGGAGG CTCCGGACCA 120  
 AGCCCCTTCA GCTTCTCCCT CCGGATCGAT GTGCTGCTGT TAACCCGTGA GGAGGCGGCG 180  
 GCGGCGGCAG CGGCAGCGGA AGATGGTGTT GCTGAGAGTG TTAATTCTGC TCCTCTCCTG 240  
 40 GGCGGCGGGG ATGGGAGGTC AGTATGGGAA TCCTTTAAAT AAATATATCA GACATTATGA 300  
 AGGATTATCT TACAATGTGG ATTCATTACA CCAAAAACAC CAGCGTGCCA AAAGAGCAGT 360  
 CTCACATGAA GACCAATTTT TACGTCTAGA TTTCCATGCC CATGGAAGAC ATTTCAACCT 420  
 ACGAATGAAG AGGGACACTT CCCTTTTCAG TGATGAATT AAAGTAGAAA CATCAAATAA 480  
 AGTACTTGAT TATGATACCT CTCATATTTA CACTGGACAT ATTTATGGTG AAGAAGGAAG 540  
 45 TTTAGCCATG GGTCTGTTAT TGATGGAAGA TTTGAAGGAT TCATCCAGAC TCGTGGTGGC 600  
 ACATTTTATG TTTGAGCCAG CAGAGAGATA TATTAAAGAC CGAACTCTGC CATTTCACTC 660  
 TGTCATTAT CATGAAGATG ATATTAATA TCCCCATAAA TACGGTCCTC AGGGGGGCTG 720  
 TGCAGATCAT TCAGTATTTG AAAGAATGAG GAAATACCAG ATGACTGGTG TAGAGGAAGT 780  
 AACACAGATA CCTCAAGAAG AACATGCTGC TAATGGTCCA GAACTTCTGA GGAAAAAACG 840  
 50 TACAAATTCA GCTGAAAAAA ATACTTGTC GCTTTATATT CAGACTGATC ATTTGTTCTT 900  
 TAAATATTAC GGAACACGAG AAGCTGTGAT TGCCAGATA TCCAGTCATG TTAAAGCGAT 960  
 TGATACAATT TACCAGACCA CAGACTTCTC CGGAATCCGT AACATCAGTT TCATGGTGAA 1020  
 ACGCATAAGA ATCAATACAA CTGCTGATGA GAAGGACCCT ACAAATCCTT TCCGTTTCCC 1080  
 AAATATTGGT GTGGAGAAGT TTCTGGAATT GAATTCTGAG CAGAATCATG ATGACTACTG 1140  
 55 TTTGGCCTAT GTCTTCACAG ACCGAGATTT TGATGATGGC GTACTTGGTC TGGCTTGGGT 1200

5

10

15

20

25

30

35

40

45

50

55

TGGAGCACCT TCAGGAAGCT CTGGAGGAAT ATGTGAAAAA AGTAAACTCT ATTCAGATGG 1260  
 TAAGAAGAAG TCCTTAAACA CTGGAATTAT TACTGTTCAG AACTATGGGT CTCATGTACC 1320  
 TCCCAAAGTC TCTCACATTA CTTTGTCTCA CGAAGTTGGA CATAACTTTG GATCCCCACA 1380  
 TGATTCTGGA ACAGAGTGCA CACCAGGAGA ATCTAAGAAT TTGGGTCAAA AAGAAAATGG 1440  
 CAATTACATC ATGTATGCAA GAGCAACATC TGGGGACAAA CTTAACAACA ATAAATTCTC 1500  
 ACTCTGTAGT ATTAGAAATA TAAGCCAAGT TCTTGAGAAG AAGAGAAAACA ACTGTTTTGT 1560  
 TGAATCTGGC CAACCTATTT GTGGAAATGG AATGGTAGAA CAAGGTGAAG AATGTGATTG 1620  
 TGGCTATAGT GACCAGTGTA AAGATGAATG CTGCTTCGAT GCAAATCAAC CAGAGGGAAG 1680  
 AAAATGCAAA CTGAAACCTG GGAAACAGTG CAGTCCAAGT CAAGGTCCTT GTTGTACAGC 1740  
 ACAGTGTGCA TTCAAGTCAA AGTCTGAGAA GTGTCGGGAT GATTCAGACT GTGCAAGGGA 1800  
 AGGAATATGT AATGGCTTCA CAGCTCTCTG CCCAGCATCT GACCCTAAAC CAAACTTCAC 1860  
 AGACTGTAAT AGGCATACAC AAGTGTGCAT TAATGGGCAA TGTGCAGGTT CTATCTGTGA 1920  
 GAAATATGGC TTAGAGGAGT GTACGTGTGC CAGTTCTGAT GGCAAAGATG ATAAAGAATT 1980  
 ATGCCATGTA TGCTGTATGA AGAAAATGGA CCCATCAACT TGTGCCAGTA CAGGGTCTGT 2040  
 GCAGTGGAGT AGGCACTTCA GTGGTCAAC CATCACCTG CAACCTGGAT CCCCTTGCAA 2100  
 CGATTTTAGA GGTACTGTG ATGTTTTTAT GCGGTGCAGA TTAGTAGATG CTGATGGTCC 2160  
 TCTAGCTAGG CTTAAAAAG CAATTTTTAG TCCAGAGCTC TATGAAAACA TTGCTGAATG 2220  
 GATTGTGGCT CATTTGGTGGG CAGTATTACT TATGGGAATT GCTCTGATCA TGCTAATGGC 2280  
 TGGATTTATT AAGATATGCA GTGTTTCATC TCCAAGTAGT AATCCAAAGT TGCCTCCTCC 2340  
 TAAACCACTT CCAGGCACTT TAAAGAGGAG GAGACCTCCA CAGCCCATTG AGCAACCCCA 2400  
 GCGTCACGGC CCCCAGAGAGA GTTATCAAAT GGGACACATG AGACGCTAAC TGCAGCTTTT 2460  
 GCCTTGGTTC TTCCTAGTGC CTACAATGGG AAAACTTCAC TCCAAAGAGA AACCTATTAA 2520  
 GTCATCATCT CCAAACTAAA CCCTCACAAAG TAACAGTTGA AGAAAAAATG GCAAGAGATC 2580  
 ATATCCTCAG ACCAGGTGGA ATTACTTAAA TTTTAAAGCC TGAAAATTCC AATTTGGGGG 2640  
 TGGGAGGTGG AAAAGGAACC CAATTTTCTT ATGAACAGAT ATTTTAACT TAATGGCACA 2700  
 AAGTCTTAGA ATATTATTAT GTGCCCCGTG TTCCCTGTTC TTCGTTGCTG CATTTTCTTC 2760  
 ACTTGCAGGC AAACCTGGCT CTCAATAAAC TTTTTCG 2796

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 748 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Val Leu Leu Arg Val Leu Ile Leu Leu Ser Trp Ala Ala Gly  
 1 5 10 15  
 Met Gly Gly Gln Tyr Gly Asn Pro Leu Asn Lys Tyr Ile Arg His Tyr  
 20 25 30  
 Glu Gly Leu Ser Tyr Asn Val Asp Ser Leu His Gln Lys His Gln Arg  
 35 40 45  
 Ala Lys Arg Ala Val Ser His Glu Asp Gln Phe Leu Arg Leu Asp Phe  
 50 55 60  
 His Ala His Gly Arg His Phe Asn Leu Arg Met Lys Arg Asp Thr Ser  
 65 70 75 80  
 Leu Phe Ser Asp Glu Phe Lys Val Glu Thr Ser Asn Lys Val Leu Asp  
 85 90 95  
 Tyr Asp Thr Ser His Ile Tyr Thr Gly His Ile Tyr Gly Glu Glu Gly  
 100 105 110  
 Ser Leu Ala Met Gly Leu Leu Leu Met Glu Asp Leu Lys Asp Ser Ser  
 115 120 125  
 Arg Leu Val Val Ala His Phe Met Phe Glu Pro Ala Glu Arg Tyr Ile  
 130 135 140  
 Lys Asp Arg Thr Leu Pro Phe His Ser Val Ile Tyr His Glu Asp Asp

145 150 155 160  
Ile Asn Tyr Pro His Lys Tyr Gly Pro Gln Gly Gly Cys Ala Asp His  
165 170 175  
5 Ser Val Phe Glu Arg Met Arg Lys Tyr Gln Met Thr Gly Val Glu Glu  
180 185 190  
Val Thr Gln Ile Pro Gln Glu Glu His Ala Ala Asn Gly Pro Glu Leu  
195 200 205  
Leu Arg Lys Lys Arg Thr Asn Ser Ala Glu Lys Asn Thr Cys Gln Leu  
210 215 220  
10 Tyr Ile Gln Thr Asp His Leu Phe Phe Lys Tyr Tyr Gly Thr Arg Glu  
225 230 235 240  
Ala Val Ile Ala Gln Ile Ser Ser His Val Lys Ala Ile Asp Thr Ile  
245 250 255  
15 Tyr Gln Thr Thr Asp Phe Ser Gly Ile Arg Asn Ile Ser Phe Met Val  
260 265 270  
Lys Arg Ile Arg Ile Asn Thr Thr Ala Asp Glu Lys Asp Pro Thr Asn  
275 280 285  
Pro Phe Arg Phe Pro Asn Ile Gly Val Glu Lys Phe Leu Glu Leu Asn  
290 295 300  
20 Ser Glu Gln Asn His Asp Asp Tyr Cys Leu Ala Tyr Val Phe Thr Asp  
305 310 315 320  
Arg Asp Phe Asp Asp Gly Val Leu Gly Leu Ala Trp Val Gly Ala Pro  
325 330 335  
25 Ser Gly Ser Ser Gly Gly Ile Cys Glu Lys Ser Lys Leu Tyr Ser Asp  
340 345 350  
Gly Lys Lys Lys Ser Leu Asn Thr Gly Ile Ile Thr Val Gln Asn Tyr  
355 360 365  
Gly Ser His Val Pro Pro Lys Val Ser His Ile Thr Phe Ala His Glu  
370 375 380  
30 Val Gly His Asn Phe Gly Ser Pro His Asp Ser Gly Thr Glu Cys Thr  
385 390 395 400  
Pro Gly Glu Ser Lys Asn Leu Gly Gln Lys Glu Asn Gly Asn Tyr Ile  
405 410 415  
35 Met Tyr Ala Arg Ala Thr Ser Gly Asp Lys Leu Asn Asn Asn Lys Phe  
420 425 430  
Ser Leu Cys Ser Ile Arg Asn Ile Ser Gln Val Leu Glu Lys Lys Arg  
435 440 445  
Asn Asn Cys Phe Val Glu Ser Gly Gln Pro Ile Cys Gly Asn Gly Met  
450 455 460  
40 Val Glu Gln Gly Glu Glu Cys Asp Cys Gly Tyr Ser Asp Gln Cys Lys  
465 470 475 480  
Asp Glu Cys Cys Phe Asp Ala Asn Gln Pro Glu Gly Arg Lys Cys Lys  
485 490 495  
45 Leu Lys Pro Gly Lys Gln Cys Ser Pro Ser Gln Gly Pro Cys Cys Thr  
500 505 510  
Ala Gln Cys Ala Phe Lys Ser Lys Ser Glu Lys Cys Arg Asp Asp Ser  
515 520 525  
Asp Cys Ala Arg Glu Gly Ile Cys Asn Gly Phe Thr Ala Leu Cys Pro  
530 535 540  
50 Ala Ser Asp Pro Lys Pro Asn Phe Thr Asp Cys Asn Arg His Thr Gln  
545 550 555 560  
Val Cys Ile Asn Gly Gln Cys Ala Gly Ser Ile Cys Glu Lys Tyr Gly  
565 570 575  
55 Leu Glu Glu Cys Thr Cys Ala Ser Ser Asp Gly Lys Asp Asp Lys Glu  
580 585 590



Leu Cys His Val Cys Cys Met Lys Lys Met Asp Pro Ser Thr Cys Ala  
 595 600 605  
 Ser Thr Gly Ser Val Gln Trp Ser Arg His Phe Ser Gly Arg Thr Ile  
 610 615 620  
 Thr Leu Gln Pro Gly Ser Pro Cys Asn Asp Phe Arg Gly Tyr Cys Asp  
 625 630 635 640  
 Val Phe Met Arg Cys Arg Leu Val Asp Ala Asp Gly Pro Leu Ala Arg  
 645 650 655  
 Leu Lys Lys Ala Ile Phe Ser Pro Glu Leu Tyr Glu Asn Ile Ala Glu  
 660 665 670  
 Trp Ile Val Ala His Trp Trp Ala Val Leu Leu Met Gly Ile Ala Leu  
 675 680 685  
 Ile Met Leu Met Ala Gly Phe Ile Lys Ile Cys Ser Val His Thr Pro  
 690 695 700  
 Ser Ser Asn Pro Lys Leu Pro Pro Pro Lys Pro Leu Pro Gly Thr Leu  
 705 710 715 720  
 Lys Arg Arg Arg Pro Pro Gln Pro Ile Gln Gln Pro Gln Arg Gln Arg  
 725 730 735  
 Pro Arg Glu Ser Tyr Gln Met Gly His Met Arg Arg  
 740 745

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2098 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GAATTCTGAG	CAGAATCATG	ATGACTACTG	TTTGGCCTAT	GTCTTCACAG	ACCGAGATTT	60
TGATGATGGC	GTACTTGGTC	TGGCTTGGGT	TGGAGCACCT	TCAGGAAGCT	CTGGAGGAAT	120
ATGTGAAAAA	AGTAACTCT	ATTGAGATGG	TAAGAAGAAG	TCCTTAAACA	CTGGAATTAT	180
TACTGTTTCA	AACTATGGGT	CTCATGTACC	TCCCAAAGTC	TCTCACATTA	CTTTTGCTCA	240
CGAAGTTGGA	CATAACTTTG	GATCCCCACA	TGATTCTGGA	ACAGAGTGCA	CACCAGGAGA	300
ATCTAAGAAT	TTGGGTCAAA	AAGAAAATGG	CAATTACATC	ATGTATGCAA	GAGCAACATC	360
TGGGGACAAA	CTTAACAACA	ATAAATTCTC	ACTCTGTAGT	ATTAGAAATA	TAAGCCAAGT	420
TCTTGAGAAG	AAGAGAAACA	ACTGTTTTGT	TGAATCTGGC	CAACCTATTT	GTGGAAATGG	480
AATGGTAGAA	CAAGGTGAAG	AATGTCATTG	TGGCTATAGT	GACCAGTGTA	AAGATGAATG	540
CTGCTTCGAT	GCAAATCAAC	CAGAGGGAAG	AAAATGCAAA	CTGAAACCTG	GGAAACAGTG	600
CAGTCCAAGT	CAAGGTCCTT	GTTGTACAGC	ACAGTGTGCA	TTCAAGTCAA	AGTCTGAGAA	660
GTGTCGGGAT	GATTGAGACT	GTGCAAGGGA	AGGAATATGT	AATGGCTTCA	CAGCTCTCTG	720
CCCAGCATCT	GACCCTAAAC	CAAACCTTCA	AGACTGTAAT	AGGCATACAC	AAGTGTGCAT	780
TAATGGGGTA	AGCATTTAAC	TATATGTTTT	AAAATTTAAT	TTTAGAAAAC	TTGTTTTTCA	840
GAAGAATTAT	TGATGCTTAA	AGCTACATAG	TTAAAGTAAT	TAATCTTGGT	CTCTGTTTAA	900
GTAATATTCC	CTCACAAAAC	CATGAATATA	TTATGTGGCA	TTCAATTAGC	TACTAATTTG	960
TCTTTCATCT	TTCCATGTAC	ATGTGGTTGA	TATTCTCTAG	AGAAACATAG	TTGTACAACT	1020
CGGCATGTGA	TTTGTCTATA	ATATTTAAGT	TTTATAAAAT	AATATTTTCA	TAGCCTAAAT	1080
AAAAGAACTC	TTTGGTCATC	TTCTCTGAAT	ATCAAACCTT	CAAAGCTTTT	GTGGCTGAAT	1140
ATCACTTTGC	TCTACAGGAA	AAAAATTTAA	TTTTTCTTTC	TTTATAGAAG	AGCCGTAATA	1200
ACCAACATAA	AATCGATCCT	CATCTAATCT	CTTGCTCTGC	TTTTATTTC	TTTTTTTAAG	1260
TTGCCATTGC	TTTAAAAGAT	TTACTATCTT	TCTTGGATTT	ACTGTTTTTC	AAATTTTTTC	1320
AAATGTATTT	ATGTAATTCA	GTTTTGATAC	TCATCTCTGT	TTGTTTTTCA	CTTTCATTTT	1380
CATTTAAATA	TTTTGACATT	GGAAGCTCAT	ACTTGCCTGT	CTGTTACTAT	AAAAAATAGG	1440
TTTGACTGTA	TAGGGATTAA	ACAATTTGTC	TTTTATTTTC	TTCTAGCAAT	GTGCAGGTTT	1500
TATCTGTGAG	AAATATGGCT	TAGAAGAGTG	TACGTGTGCC	AGTCTGATGG	CAAAGATGAT	1560

AAAGAATTAT	GCCATGTATG	CTGTATGAAG	AAAAGTAAGG	CTTTTAAAAA	CACAAGATAT	1620
AAAATTTGCC	TCAAACATTT	ATTTTCTCCT	AAATTTTAAAG	TGTAAACTT	TGACCTACAG	1680
TTTGCCGAGA	TAATTTCCAG	CTAAATCTGT	CCTCTTGAGG	AGATTATAAA	TGTAACGTAG	1740
CATTGTGTCT	CTATTATTAT	GGTCTCTACA	ATGTTTTTAA	AATGATAAAC	TAGACAAAAC	1800
GTTGCCAGCT	TTACAGCAGT	AATTTACATA	AACACTGTTA	GACTTTAAGT	CATCGTGGAC	1860
ACTGAGTCAA	GACTTGCTGG	TTGCTTGTTT	ACATTGTAAC	ATTTAATATG	AATTACTGAT	1920
GGCGTTACCC	AGCCTAACTA	GAGAAGGTCT	GTATAACATG	TTATGGTAAT	GATTTTCAGTT	1980
TTTTTTCCCT	CTTTGTATTT	GCACAACCTG	GAAATCTGAT	CTGCAACTTA	TATTTGAATC	2040
TGACCTTCAG	CTTATATTTG	GCATTTCTTT	TCCAGTGGAC	CCATCAACTC	CGGAATTC	2098

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 265 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asn	Ser	Glu	Gln	Asn	His	Asp	Asp	Tyr	Cys	Leu	Ala	Tyr	Val	Phe	Thr
1				5					10					15	
Asp	Arg	Asp	Phe	Asp	Asp	Gly	Val	Leu	Gly	Leu	Ala	Trp	Val	Gly	Ala
			20					25					30		
Pro	Ser	Gly	Ser	Ser	Gly	Gly	Ile	Cys	Glu	Lys	Ser	Lys	Leu	Tyr	Ser
		35				40					45				
Asp	Gly	Lys	Lys	Lys	Ser	Leu	Asn	Thr	Gly	Ile	Ile	Thr	Val	Gln	Asn
	50					55					60				
Tyr	Gly	Ser	His	Val	Pro	Pro	Lys	Val	Ser	His	Ile	Thr	Phe	Ala	His
	65				70					75				80	
Glu	Val	Gly	His	Asn	Phe	Gly	Ser	Pro	His	Asp	Ser	Gly	Thr	Glu	Cys
			85					90						95	
Thr	Pro	Gly	Glu	Ser	Lys	Asn	Leu	Gly	Gln	Lys	Glu	Asn	Gly	Asn	Tyr
			100					105					110		
Ile	Met	Tyr	Ala	Arg	Ala	Thr	Ser	Gly	Asp	Lys	Leu	Asn	Asn	Asn	Lys
		115						120					125		
Phe	Ser	Leu	Cys	Ser	Ile	Arg	Asn	Ile	Ser	Gln	Val	Leu	Glu	Lys	Lys
		130				135					140				
Arg	Asn	Asn	Cys	Phe	Val	Glu	Ser	Gly	Gln	Pro	Ile	Cys	Gly	Asn	Gly
					150					155				160	
Met	Val	Glu	Gln	Gly	Glu	Glu	Cys	Asp	Cys	Gly	Tyr	Ser	Asp	Gln	Cys
			165					170						175	
Lys	Asp	Glu	Cys	Cys	Phe	Asp	Ala	Asn	Gln	Pro	Glu	Gly	Arg	Lys	Cys
			180					185					190		
Lys	Leu	Lys	Pro	Gly	Lys	Gln	Cys	Ser	Pro	Ser	Gln	Gly	Pro	Cys	Cys
		195					200					205			
Thr	Ala	Gln	Cys	Ala	Phe	Lys	Ser	Lys	Ser	Glu	Lys	Cys	Arg	Asp	Asp
	210					215					220				
Ser	Asp	Cys	Ala	Arg	Glu	Gly	Ile	Cys	Asn	Gly	Phe	Thr	Ala	Leu	Cys
	225				230					235				240	
Pro	Ala	Ser	Asp	Pro	Lys	Pro	Asn	Phe	Thr	Asp	Cys	Asn	Arg	His	Thr
			245					250						255	
Gln	Val	Cys	Ile	Asn	Gly	Val	Ser	Ile							
			260					265							

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2481 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCGTGAGGAG	GCGGCGGCCG	GGAAGATGGT	GTTGCCGACA	GTGTTAATTC	TGCTCCTCTC	60
CTGGGCGGCG	GGGCTGGGAG	GTCTAGTATG	AAATCCTTTA	AATAAATATA	TTAGACATTA	120
TGAAGGATTA	TCTTACAATG	TGGATTCATT	ACACCAAAAA	CACCAGCGTG	CCAAACGAGC	180
AGTCTCACAT	GAGGACCAGT	TTTTACTTCT	AGATTTCCAT	GCTCATGGAA	GACAGTTCAA	240
CCTACGAATG	AAGAGGGACA	CTTCCCTTTT	TAGTGATGAA	TTTAAAGTAG	AAACATCAAA	300
TAAAGTACTT	GATTATGATA	CCTCTCATAT	TTACACTGGA	CATATTTATG	GTGAAGAAGG	360
AAGCTTTAGT	CATGGGTCTG	TCATTGATGG	AAGATTTGAA	GGTTTCATCA	AGACTCGTGG	420
TGGCACGTTT	TACATTGAGC	CAGCAGAGAG	ATACATTAAA	GATCGAATCC	TGCCATTTC	480
CTCTGTCAAT	TATCATGAAG	ATGATATTAA	CTATCCCCAT	AAATACGGCC	CACAGGGGGG	540
CTGTGCAGAT	CACTCCGTTT	TTGAAAGGAT	GAGGAAGTAC	CAAATGACTG	GAGTAGAGGA	600
AGGAGCCCCG	GCACATCCAG	AGAAGCATGC	TGCTAGTAGT	GGTCCTGAGC	TCCTGAGGAA	660
AAAACGCACA	ACTCTGGCTG	AAAGAAATAC	TTGTCAGCTC	TATATCCAGA	CAGATCACCT	720
GTTCTTTAAA	TACTATGGAA	CACGAGAAGC	TGTGATTGCT	CAGATATCCA	GTCATGTTAA	780
AGCAATTGAT	ACAATTTACC	AGACTACAGA	CTTCTCCGGA	ATCCGTAACA	TCAGCTTCAT	840
GGTGAAACGC	ATAAGAATCA	ATACAACCTC	TGATGAAAAA	GACCCTACAA	ATCCTTTCCG	900
TTTCCCAAAT	ATTGGTGTGG	AGAAGTTTCT	GGAGTTGAAT	TCTGAGCAGA	ATCATGATGA	960
CTACTGCCTG	GCCTATGTCT	TCACAGACCG	GGATTTTGAT	GATGGTGTTC	TTGGTCTGGC	1020
CTGGGTGGA	GCACCTTCAG	GAAGCTCTGG	GGGAATATGT	GAGAAAAGCA	AGTTGTATTC	1080
AGATGGCAAG	AAGAAGTCAT	TGAACACAGG	CATCATTACT	GTTTCAGAACT	ATGGCTCCCA	1140
TGTGCCTCCC	AAAGTCTCTC	ATATTACGTT	TGCTCATGAA	GTTGGACATA	ACTTTGGATC	1200
TCCACATGAT	TCTGGAACAG	AGTGTAATCC	AGGAGAGTCT	AAGAACTTAG	GACAAAAAGA	1260
AAATGGCAAT	TACATCATGT	ATGCAAGAGC	AACATCTGGG	GACAACTTA	ACAACAACAA	1320
ATTTTCACTC	TGCAGCATT	GAAACATAAG	CCAAGTGCTT	GAGAAGAAGA	GGAACAACTG	1380
TTTGTGTGAA	TCTGGCCAGC	CTATCTGTGG	AAACGGGATG	GTGGAACAAG	GGGAAGAGTG	1440
TGACTGTGGC	TACAGTGACC	AGTGCAAGA	TGATTGCTGC	TTCGATGCCA	ACCAGCCAGA	1500
GGGGAAGAAA	TGCAAGCTGA	AGCCTGGGAA	GCAGTGCAAG	CCGAGTCAAG	GACCCTGCTG	1560
TACAGCACAG	TGTGCATTCA	AGTCAAAGTC	TGAAAAGTGC	CGGGATGATT	CTGACTGTGC	1620
AAAGGAAGGG	ATATGCAATG	GCTTCACAGC	CCTTTGCCCA	GCATCTGATC	CCAAGCCCAA	1680
CTTTACAGAC	TGTAACAGGC	ACACACAAGT	GTGCATTAAT	GGGCAATGTG	CAGGTTCTAT	1740
TTGTGAAAAG	TATGACTTGG	AGGAGTGCAC	CTGTGCCAGC	TCTGATGGCA	AAGATAATAA	1800
GGAATTATGC	CATGTTTGTG	GCATGAAGAA	AATGGCTCCA	TCAACTTGTG	CCAGTACAGG	1860
CTCTTTGCAG	TGGAGCAAGC	AGTTCACTGG	TCGGACTATC	ACTCTGCAGC	CGGGCTCTCC	1920
ATGTAATGAC	TTCAGAGGCT	ACTGTGATGT	TTTCATGCGG	TGCAGATTAG	TAGATGCTGA	1980
TGGCCCTCTA	GCTAGGCTGA	AAAAAGCCAT	TTTTAGTCCA	CAACTCTATG	AAAACATTGC	2040
TGAGTGGATT	GTGGCTCACT	GGTGGGCAGT	ACTGCTTATG	GGAATTGCCC	TGATCATGTT	2100
AATGGCTGGA	TTTATCAAGA	TTTGCAGTGT	TCACACTCCA	AGTAGTAATC	CAAAGTTGCC	2160
GCCTCCTAAA	CCACTTCCAG	GCACTTTAAA	GAGGAGGAGA	CCGCCACAGC	CCATTTCAGCA	2220
GCCCCGCGT	CAGAGGCCCC	GAGAGAGTTA	TCAAATGGGA	CACATGCGAC	GCTAATGCAG	2280
CTTTTGCCCT	GGTTCCTTCT	AGTGCCTACA	GTGGGAAAAC	TTCACTCCAA	AGAGAAACCT	2340
GTTAAGTCAT	CATCTGCAAA	TGATACCCTT	ACAGTTAATA	GTTGAAGAAA	AAATGGCAAG	2400
AGATCATGTC	CTCAGATCAG	GTGGAATTAC	TCAAAATTTA	AAGCCTGAAA	ATTCCAATTT	2460
TGGGGGTGGG	GGTGGGATGG	G				2481

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 749 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Val Leu Pro Thr Val Leu Ile Leu Leu Leu Ser Trp Ala Ala Gly  
1 5 10 15  
Leu Gly Gly Gln Tyr Gly Asn Pro Leu Asn Lys Tyr Ile Arg His Tyr  
20 25 30  
Glu Gly Leu Ser Tyr Asn Val Asp Ser Leu His Gln Lys His Gln Arg  
35 40 45  
Ala Lys Arg Ala Val Ser His Glu Asp Gln Phe Leu Leu Leu Asp Phe  
50 55 60  
His Ala His Gly Arg Gln Phe Asn Leu Arg Met Lys Arg Asp Thr Ser  
65 70 75 80  
Leu Phe Ser Asp Glu Phe Lys Val Glu Thr Ser Asn Lys Val Leu Asp  
85 90 95  
Tyr Asp Thr Ser His Ile Tyr Thr Gly His Ile Tyr Gly Glu Glu Gly  
100 105 110  
Ser Phe Ser His Gly Ser Val Ile Asp Gly Arg Phe Glu Gly Phe Ile  
115 120 125  
Lys Thr Arg Gly Gly Thr Phe Tyr Ile Glu Pro Ala Glu Arg Tyr Ile  
130 135 140  
Lys Asp Arg Ile Leu Pro Phe His Ser Val Ile Tyr His Glu Asp Asp  
145 150 155 160  
Ile Asn Tyr Pro His Lys Tyr Gly Pro Gln Gly Gly Cys Ala Asp His  
165 170 175  
Ser Val Phe Glu Arg Met Arg Lys Tyr Gln Met Thr Gly Val Glu Glu  
180 185 190  
Gly Ala Arg Ala His Pro Glu Lys His Ala Ala Ser Ser Gly Pro Glu  
195 200 205  
Leu Leu Arg Lys Lys Arg Thr Leu Ala Glu Arg Asn Thr Cys Gln  
210 215 220  
Leu Tyr Ile Gln Thr Asp His Leu Phe Phe Lys Tyr Tyr Gly Thr Arg  
225 230 235 240  
Glu Ala Val Ile Ala Gln Ile Ser Ser His Val Lys Ala Ile Asp Thr  
245 250 255  
Ile Tyr Gln Thr Thr Asp Phe Ser Gly Ile Arg Asn Ile Ser Phe Met  
260 265 270  
Val Lys Arg Ile Arg Ile Asn Thr Thr Ser Asp Glu Lys Asp Pro Thr  
275 280 285  
Asn Pro Phe Arg Phe Pro Asn Ile Gly Val Glu Lys Phe Leu Glu Leu  
290 295 300  
Asn Ser Glu Gln Asn His Asp Asp Tyr Cys Leu Ala Tyr Val Phe Thr  
305 310 315 320  
Asp Arg Asp Phe Asp Asp Gly Val Leu Gly Leu Ala Trp Val Gly Ala  
325 330 335  
Pro Ser Gly Ser Ser Gly Gly Ile Cys Glu Lys Ser Lys Leu Tyr Ser  
340 345 350  
Asp Gly Lys Lys Lys Ser Leu Asn Thr Gly Ile Ile Thr Val Gln Asn  
355 360 365  
Tyr Gly Ser His Val Pro Pro Lys Val Ser His Ile Thr Phe Ala His  
370 375 380  
Glu Val Gly His Asn Phe Gly Ser Pro His Asp Ser Gly Thr Glu Cys  
385 390 395 400  
Thr Pro Gly Glu Ser Lys Asn Leu Gly Gln Lys Glu Asn Gly Asn Tyr  
405 410 415  
Ile Met Tyr Ala Arg Ala Thr Ser Gly Asp Lys Leu Asn Asn Asn Lys

5

10

15

20

25

30

35

40

45

			420					425				430			
Phe	Ser	Leu	Cys	Ser	Ile	Arg	Asn	Ile	Ser	Gln	Val	Leu	Glu	Lys	Lys
			435					440				445			
Arg	Asn	Asn	Cys	Phe	Val	Glu	Ser	Gly	Gln	Pro	Ile	Cys	Gly	Asn	Gly
			450					455				460			
Met	Val	Glu	Gln	Gly	Glu	Glu	Cys	Asp	Cys	Gly	Tyr	Ser	Asp	Gln	Cys
465					470					475				480	
Lys	Asp	Asp	Cys	Cys	Phe	Asp	Ala	Asn	Gln	Pro	Glu	Gly	Lys	Lys	Cys
				485					490					495	
Lys	Leu	Lys	Pro	Gly	Lys	Gln	Cys	Ser	Pro	Ser	Gln	Gly	Pro	Cys	Cys
			500					505					510		
Thr	Ala	Gln	Cys	Ala	Phe	Lys	Ser	Lys	Ser	Glu	Lys	Cys	Arg	Asp	Asp
			515					520					525		
Ser	Asp	Cys	Ala	Lys	Glu	Gly	Ile	Cys	Asn	Gly	Phe	Thr	Ala	Leu	Cys
	530					535					540				
Pro	Ala	Ser	Asp	Pro	Lys	Pro	Asn	Phe	Thr	Asp	Cys	Asn	Arg	His	Thr
545					550					555				560	
Gln	Val	Cys	Ile	Asn	Gly	Gln	Cys	Ala	Gly	Ser	Ile	Cys	Glu	Lys	Tyr
				565					570					575	
Asp	Leu	Glu	Glu	Cys	Thr	Cys	Ala	Ser	Ser	Asp	Gly	Lys	Asp	Asn	Lys
			580					585					590		
Glu	Leu	Cys	His	Val	Cys	Cys	Met	Lys	Lys	Met	Ala	Pro	Ser	Thr	Cys
			595				600					605			
Ala	Ser	Thr	Gly	Ser	Leu	Gln	Trp	Ser	Lys	Gln	Phe	Ser	Gly	Arg	Thr
	610					615					620				
Ile	Thr	Leu	Gln	Pro	Gly	Ser	Pro	Cys	Asn	Asp	Phe	Arg	Gly	Tyr	Cys
625					630					635				640	
Asp	Val	Phe	Met	Arg	Cys	Arg	Leu	Val	Asp	Ala	Asp	Gly	Pro	Leu	Ala
				645					650					655	
Arg	Leu	Lys	Lys	Ala	Ile	Phe	Ser	Pro	Gln	Leu	Tyr	Glu	Asn	Ile	Ala
			660					665						670	
Glu	Trp	Ile	Val	Ala	His	Trp	Trp	Ala	Val	Leu	Leu	Met	Gly	Ile	Ala
			675					680					685		
Leu	Ile	Met	Leu	Met	Ala	Gly	Phe	Ile	Lys	Ile	Cys	Ser	Val	His	Thr
	690					695					700				
Pro	Ser	Ser	Asn	Pro	Lys	Leu	Pro	Pro	Pro	Lys	Pro	Leu	Pro	Gly	Thr
705					710					715				720	
Leu	Lys	Arg	Arg	Arg	Pro	Pro	Gln	Pro	Ile	Gln	Gln	Pro	Pro	Arg	Gln
				725					730					735	
Arg	Pro	Arg	Glu	Ser	Tyr	Gln	Met	Gly	His	Met	Arg	Arg			
			740					745							